DOCUMENT RESUME

ED 447 201 TM 032 090

AUTHOR McMillan, James H.
TITLE Basic Assessment Concepts for Teachers and School Administrators. ERIC/AE Digest.
INSTITUTION ERIC Clearinghouse on Assessment and Evaluation, College Park, MD.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO EDO-TM-00-09
PUB DATE 2000-11-00
NOTE 4p.; Version of a paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
CONTRACT ED-99-CO-0032
PUB TYPE ERIC Publications (071) -- ERIC Digests in Full Text (073) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Administrators; Educational Assessment; *Educational Testing; Elementary Secondary Education; *Student Evaluation; *Teachers; Test Construction; *Test Use
IDENTIFIERS ERIC Digests

ABSTRACT

This digest presents 11 basic principles to guide the assessment training and professional development of teachers and administrators. These are: (1) assessment is inherently a process of professional judgment; (2) assessment is based on separate but related principles of measurement evidence and evaluation; (3) assessment decision-making is influenced by a series of tensions; (4) assessment influences student motivation and learning; (5) assessment contains error; (6) good assessment enhances instruction; (7) good assessment is valid; (8) good assessment is fair and ethical; (9) good assessments use multiple methods; (10) good assessment is efficient and feasible; and (11) good assessment appropriately incorporates technology. What is most essential about assessment is understanding how general, fundamental assessment principles and ideas can be used to enhance student learning and teacher effectiveness. (Contains 13 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 447 201

Basic Assessment Concepts for Teachers and School Administrators

James H. McMillan
Virginia Commonwealth University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM032090

BEST COPY AVAILABLE

ERIC Digest EDO-TM-00-09 November 2000

ERIC®/AE Digest Series EDO-TM-00-09
University of Maryland, College Park

November 2000
Department of Measurement, Statistics and Evaluation

Basic Assessment Concepts for Teachers and School Administrators

James H. McMillan
Virginia Commonwealth University

While several authors have argued that there are a number of essential assessment concepts, principles, techniques, and procedures that teachers and administrators need to know (e.g., Calfee & Masuda, 1997; Cizek, 1997), there continues to be relatively little emphasis on assessment in the preparation or professional development of teachers and administrators (Stiggins, 2000). Nevertheless, guidance is provided through established professional standards for assessment skills of teachers (*Standards for Teacher Competence in Educational Assessment of Students* (1990), a framework of assessment tasks for administrators (Impara & Plake, 1996), the Code of Professional Responsibilities in Educational Measurement (1995), the Code of Fair Testing Practices in Education (1988), and the new edition of *Standards for Educational and Psychological Testing* (1999).

This Digest draws on suggestions and guidelines from these sources, in light of current assessment demands and contemporary theories of learning and motivation, to present eleven basic principles to guide the assessment training and professional development of teachers and administrators.

Assessment is inherently a process of professional judgment.

Professional judgment is the foundation for all assessment. The measurement of student performance may seem "objective," given such practices as machine scoring and multiple-choice test items, but even these approaches are based on professional assumptions and values. Whether that judgment occurs in constructing test questions, scoring essays, creating rubrics, grading participation, combining scores, or interpreting standardized test scores, the essence of the process is making professional interpretations and decisions.

Assessment is based on separate but related principles of measurement evidence and evaluation.

It is important to understand the difference between measurement evidence (differentiating degrees of a trait by description or by assigning scores) and evaluation (interpreting the description or scores). Essential measurement evidence skills include the ability to understand and interpret the meaning of descriptive statistical procedures, including variability, correlation, percentiles, standard scores, growth-scale scores, norming, and principles of combining scores for grading. A conceptual understanding of these techniques is needed (not necessarily knowing how to compute statistics) for such tasks as interpreting student strengths and weaknesses, examining reliability and validity evidence, determining grades, and making admissions decisions. Schafer (1991) has indicated that these concepts and techniques comprise part of an essential language for educators. They also provide a common basis for communication about results, interpretation of evidence, and appropriate use of data. This is increasingly important given the pervasiveness of standards-based, high-stakes, large-scale assessments. Evaluation concerns merit and worth of the data as applied to a specific use or context. It involves what Shepard (2000) has described as the systematic analysis of evidence to interpret evidence effectively and make value judgments about the meaning of the results.

Assessment decision-making is influenced by a series of tensions.

Competing purposes, uses, and pressures result in tensions for teachers and administrators as they make assessment-related decisions. For example, good teaching incorporates assessments that motivate and engage students in ways that are consistent with teachers' philosophies of teaching and learning and with theories of development, learning and motivation. Most teachers want to use constructed-response assessments because they believe this kind of testing is best to ascertain student understanding. On the other hand, factors external to the classroom, such as mandated large-scale testing, promote different assessment strategies, such as using selected-response tests and providing practice in objective test-taking (McMillan & Nash, 2000). Further examples of tensions include formative (informal and ongoing) vs. summative (formal and at the end), criterion-referenced vs. norm-referenced, traditional vs. alternative, authentic vs. contrived, and standardized tests vs. classroom tests. Decisions about assessment are best made with a full understanding of how different factors influence the nature of the assessment.

Assessment influences student motivation and learning.

The nature of assessment influences what is learned and the degree of meaningful engagement by students in the learning process. While Wiggins (1998) contends that assessments should be authentic, with feedback and opportunities for revision to improve, rather than simply audit, learning, the more general principle is understanding how different assessments affect students. Will students be more engaged if assessment tasks are problem-based? How do students study when they know the test consists of multiple-choice items? What is the nature of feedback, and when is it given to students? How does assessment affect student effort? Recent research summarized by Black & Wiliam (1998), for example, shows that student self-assessment skills, learned and applied as part of formative assessment, enhance student achievement.

Assessment contains error.

Teachers and administrators need to know not only that there is error in all classroom and standardized assessments, but also, more specifically, how reliability is determined and how much error is likely. With so much emphasis today on high-stakes testing for promotion, graduation, teacher and administrator accountability, and school accreditation, it is critical that all educators understand such concepts as standard error of measurement, reliability coefficients, confidence intervals, and standard setting. Two reliability principles deserve special attention: (1) Reliability refers to scores, not instruments; (2) Typically, error is underestimated.

Good assessment enhances instruction.

Just as assessment affects student learning and motivation, it also influences the nature of instruction in the classroom. When assessment is integrated with instruction, it informs teachers about what activities and assignments will be most useful, what level of teaching is most appropriate, and how summative assessments provide diagnostic information. During instructional activities, informal, formative assessment helps teachers know when to move on, when to ask more

ERIC® ERIC Clearinghouse on Assessment and Evaluation, 1129 Shriver Laboratory,
University of Maryland, College Park, MD 20742 • 800 464-3742

questions, when to give more examples, and how to respond to students questions. Standardized test scores, when used appropriately, help teachers understand student strengths and weaknesses in order to target further instruction.

Good assessment is valid.

As with reliability, validity has certain technical terms and issues associated with it that are essential in helping teachers and administrators make reasonable and appropriate inferences from assessment results (e.g., types of validity evidence, validity generalization, construct underrepresentation, construct-irrelevant variance, and discriminant and convergent evidence). Of critical importance is the concept of evidence based on consequences, a new major validity category in the recently revised *Standards*. Both intended and unintended consequences of assessment need to be examined with appropriate evidence that supports particular arguments or points of view. Of equal importance is getting teachers and administrators to understand their role in gathering and interpreting validity evidence.

Good assessment is fair and ethical.

Arguably, the most important change in the recently published *Standards* is an entire new major section entitled "Fairness in Testing." The *Standards* presents four views of fairness: as absence of bias (e.g., offensiveness and unfair penalization), as equitable treatment, as equality in outcomes, and as opportunity to learn. It includes entire chapters on the rights and responsibilities of test takers, testing individuals of diverse linguistic backgrounds, and testing individuals with disabilities or special needs. Three additional areas are also important: (1) Student knowledge of learning targets and the nature of assessments prior to instruction (e.g., what will be tested, how it will be graded, scoring criteria, anchors, exemplars); (2) Student prerequisite knowledge and skills, including test-taking skills; and (3) Avoidance of stereotypes.

Good assessments use multiple methods.

Assessment that is fair, leading to valid inferences with a minimum of error, is a series of measures that show student understanding through multiple methods. A complete picture of what students understand and can do is put together in pieces comprised by different approaches to assessment. While testing experts and testing companies stress that important decisions should not be made on the basis of a single test score, some educators at the local level, and some politicians at the state and national levels, seem determined to violate this principle. There is a need to understand the entire range of assessment techniques and methods and their limitations.

Good assessment is efficient and feasible.

Teachers and school administrators have limited time and resources. Consideration must be given to the efficiency of different approaches to assessment, balancing needs to implement methods required to provide a full understanding with the time needed to develop and implement the methods and score results. Teacher skills and knowledge are important to consider, as are the levels of support and resources.

Good assessment appropriately incorporates technology.

As technology advances and teachers become more proficient in the use of technology, there will be increased opportunities for teachers and administrators to use computer-based techniques (e.g., item banks, electronic grading, computer-adapted testing, computer-based simulations), Internet resources, and more complex, detailed ways of reporting results. There is, however, a danger that technology will contribute to the mindless use of new resources, such as using items online developed by some companies without adequate evidence of reliability, validity, and fairness, and crunching numbers with software programs without sufficient thought about weighting, error, and averaging.

To summarize, what is most essential about assessment is understanding how general, fundamental assessment principles and ideas can be used to enhance student learning and teacher effectiveness. This will be achieved as teachers and administrators learn about conceptual and technical assessment concepts, methods, and procedures, for both large-scale and classroom assessments, and apply these fundamentals to instruction.

Note:

An earlier version of this paper was presented at the Annual Meeting of the American Educational Research Association, New Orleans, April 24, 2000.

References

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Calfee, R.C., & Masuda, W.V. (1997). Classroom assessment as inquiry. In G.D. Phye (Ed.) *Handbook of classroom assessment: Learning, achievement, and adjustment.* NY: Academic Press.

Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at a crossroads. In G.D. Phye (Ed.) *Handbook of classroom assessment: Learning, achievement, and adjustment.* NY: Academic Press.

Code of fair testing practices in education (1988). Washington, DC: Joint Committee on Testing Practices (American Psychological Association). Available: http://ericae.net/code.htm

Code of professional responsibilities in educational measurement (1995). Washington, DC: National Council on Measurement in Education. Available: http://www.unl.edu/buros/article2.html

Impara, J. C., & Plake, B. S. (1996). Professional development in student assessment for educational administrators. *Educational Measurement: Issues and Practice, 15*(2), 14-19.

McMillan, J. H., & Nash, S. (2000). *Teachers' classroom assessment and grading decision making.* Paper presented at the Annual Meeting of the National Council of Measurement in Education, New Orleans.

Schafer, W. D. (1991). Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice,* 10(1), 3-6.

Shepard, L. A. (2000). *The role of assessment in a learning culture.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans. Available http://www.aera.net/meeting/am2000/wrap/praddr01.htm

Standards for educational and psychological testing (1999). Washington, DC: American Educational Research Association, American Psychological Association, National Council on Measurement in Education.

Standards for teacher competence in educational assessment of students. (1990). American Federation of Teachers, National Council on Measurement in Education, National Education Association. Available: http://www.unl.edu/buros/article3.html

Stiggins, R. J. (2000). *Classroom assessment: A history of neglect, a future of immense potential.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance.* San Francisco: Jossey-Bass.

BEST COPY AVAILABLE

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract ED99CO0032. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education. **Permission is granted to copy and distribute this ERIC/AE Digest.**